Spatial Statistics and Image Analysis

Lecture 8

Konstantinos Konstantinou

Mathematical sciences
Chalmers University of Technology and University of Gothenburg
Gothenburg, Sweden

Lecture's contents

Todays lecture will cover

- First/second-order properties of point processes
- Summary functions
- Poisson point process
- Log-Gaussian Cox process
- Neyman-Scott processes
- Matérn inhibition processes
- Pairwise interaction point processes

Definition: Point processes

A point process N is a stochastic mechanism or rule to produce point patterns or realisations according to the distribution of the process.

A marked point process is a point process where each point x_i of the process is assigned a quantity $m(x_i)$, called a mark. Often, marks are integers or real numbers but more general marks can also be considered.

Two interpretations

- N is a counting measure. For a subset B of $\mathbb{R}^d$, $N(B)$ is the random number of points in B. It is assumed that $N(B) < \infty$ for all bounded sets B, i.e. that N is locally finite.
- N is a random set, i.e. the set of all points $x_1, x_2, \ldots$ in the process. In other words

  $$N = \{x_i\} \text{ or } N = \{x_1, x_2, \ldots\}$$

  Therefore, $x \in N$ means that the point x is in the set N. The set N can be finite or infinite. If it is finite the total number of points can be deterministic or random.

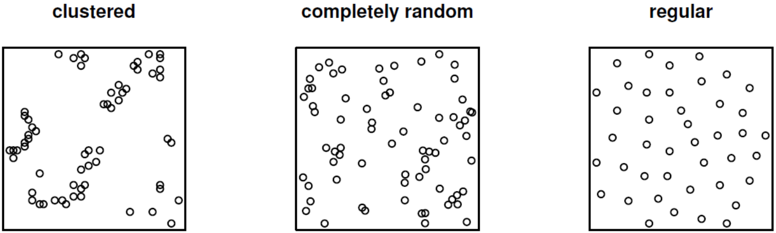
Remarks

Remark 1: We assume that all point processes are simple, i.e. that there are no multiple points ($x_i \neq x_j$ if $i \neq j$).

Remark 2: There is a large literature on processes $\{Z(t) : t \in T\}$, where T is a point process in time. There is an overlap of methods for point processes in space and in time but the temporal case is not only a special case of the spatial process with $d = 1$. Time is 1-directional.

Remark 3: To avoid confusion between points of the process and point of $\mathbb{R}^d$, the points of the process or point pattern (realization) are called events (or trees or cells).

Spatial point patterns



- One of the main questions for point pattern data is usually to determine if we have clustering or repulsion.
- The completely random case corresponds to the Poisson point process

Examples

- Locations of betacells within a rectangular region in a cat's eye (regular)
- Locations of Finnish pine saplings (clustered)
- Locations of Spanish towns (regular)
- Locations of galaxes (clustered)

Remark: Very different scales, from microscopic to cosmic

First-order properties (without marks)

The mean number of points of N in B is $\mathbb{E}(N(B))$ (depends on the set B). We use the notation

$$\Lambda(B) = \mathbb{E}(N(B))$$

and call Λ the intensity measure.

Under some continuity conditions, a density function λ, called the intensity function, exists, and

$$\Lambda(B) = \int_B \lambda(x)\, dx.$$

Some properties of point processes: stationarity and isotropy

A point process N is stationary (translation invariant) if N and the translated point process N_x have the same distribution for all translations x, i.e.

$$N = \{x_1, x_2, ...\} \text{ and } N_x = \{x_1 + x, x_2 + x, ...\}$$

have the same distribution for all $x \in \mathbb{R}^d$.

A point process is isotropic (rotation invariant) if its characteristics are invariant under rotations, i.e.

$$N = \{x_1, x_2, ...\} \text{ and } rN_x = \{rx_1, rx_2, ...\}$$

have the same distribution for any rotation r around the origin.
If a point process is both stationary and isotropic, it is called motion-invariant.

First-order properties

If N is stationary, then

$$\Lambda(B) = \lambda|B|,$$

where $0 < \lambda < \infty$ is called the intensity of N and $|B|$ is the volume of B.

λ is the mean number of points of N per unit area, i.e.

$$\lambda = \frac{\Lambda(B)}{|B|} = \frac{\mathbb{E}(N(B))}{|B|}.$$

Homogeneous Poisson point process

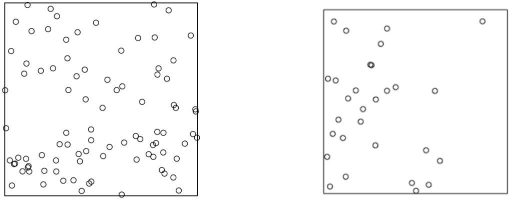
A point process is a homogeneous Poisson process (CSR) if

(P1) for some $\lambda > 0$ and any finite region B, $N(B)$ has a Poisson distribution with mean $\lambda|B|$

(P2) given $N(B) = n$, the events in B form an independent random sample from the uniform distribution on B

Inhomogeneous Poisson process: intensity λ (in homogeneous Poisson process) replaced by an intensity function $\lambda(x)$

Examples: realizations of Poisson processes



Left: Homogeneous Poisson point process with intensity 100
Right: A Poisson point process with a spatially varying intensity function $\lambda(x, y) = 100e^{-3x}$

Some comments

- Despite its simplicity the Poisson point process has a fundamental role in the characterisation of the spatial structure of spatial point processes.
- It is used as a reference model to assess clustering or inhibition.
- It is often used as a building block for constructing more complex point processes, for example, Neyman-Scott point processes(clustered) and Matérn hard-core point processes (regular).

Two distribution functions

1. Let D_1 denote the distance from an arbitrary event x to the nearest other event. Then, the nearest neighbour distance function is

$$G(r) = P(D_1 \leq r) = 1 - P(N(b(x,r) \setminus \{x\}) = 0)$$

   If the pattern is completely spatially random (CSR), $G(r) = 1 - \exp(-\lambda\pi r^2)$. For regular patterns $G(r)$ tends to lie below and for clustered patterns above the CSR curve.

2. Let D_2 denote the distance from an arbitrary point x to the nearest event. Then,

$$F(r) = P(D_2 \leq r) = 1 - P(N(b(x,r)) = 0)$$

   If the pattern is completely spatially random, $F(r) = 1 - \exp(-\lambda\pi r^2)$. For regular patterns $F(r)$ tends to lie above and for clustered patterns below the CSR curve.

Combination of the two

Using G and F we can define the so-called J function as

$$J(r) = \frac{1 - G(r)}{1 - F(r)}$$

(whenever $F(r) > 0$)

If the pattern is completely spatially random, $J(r) \equiv 1$. For regular patterns $J(r) > 1$ and for clustered patterns $J(r) < 1$.

Second-order properties

The 2nd order properties of a stationary and isotropic point process can be characterized by Ripley's K function

$K(r) = \lambda^{-1}\mathbb{E}[\#$ further events within distance r of a typical event$]$.

Often, a transformed and centered version of the K function is used,

$$L(r) - r = \sqrt{K(r)/\pi} - r,$$

which equals 0 under CSR. Values less than zero indicate regularity and values larger than zero clustering.

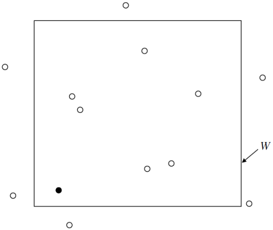
Estimation of the K function.

- An naive estimate for the K function is given by

$$\hat{K}(r) = \frac{1}{n\hat{\lambda}} \sum_{i=1}^{n} \sum_{i \neq j} 1\{\| x_i - x_j \| \leq r\}$$

where $\hat{\lambda} = \frac{n-1}{|W|}$ is an estimate for λ

- Typically, a point pattern is observed in a (bounded) observation window and points outside the window are not observed. Hence this estimator is biased.



Edge corrections

- Estimators of the summary functions need to be edge-corrected
- An unbiased estimate for the K function is given by

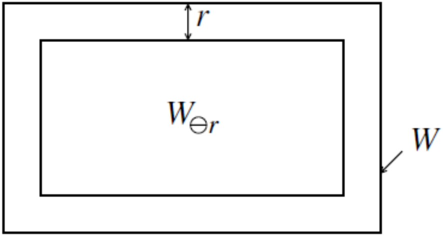
$$\hat{K}(r) = \frac{1}{n\hat{\lambda}} \sum_{i=1}^{n} \sum_{i \neq j} w(x_i, x_j) 1\{\| x_i - x_j \| \leq r\}$$

  where $w(x_i, x_j)$ is an edge correction term.
- Edge correction methods include minus sampling or border method, Ripley's isotropic correction and translation (stationary) correction

Minus sampling/ border correction.

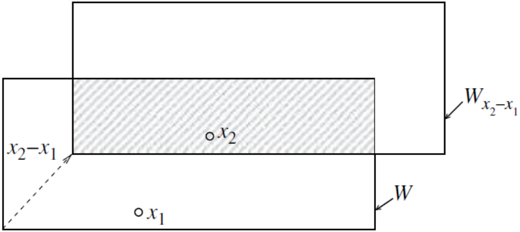
- Let $W_{\ominus r}$ denote the subset of W the points in which are in the interior of W and have a distance larger than r from the boundary ∂W. We consider pair of points $x \in W_{\ominus r}$ and $y \in W$. Then an unbiased estimate for K is given by

$$\hat{K}_{minus}(r) = \frac{1}{\hat{\lambda}^2 \mid W_{\ominus r} \mid} \sum_{x \in X \cap W_{\ominus r}} \sum_{y \in X} 1\{\| x - y \| \leq r\}$$



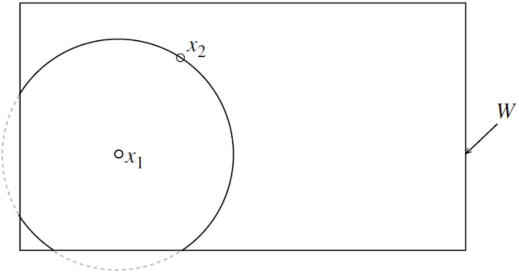
Translational edge-correction

- Applicable for stationary point processes
- The weights $w(x_i, x_j) = \frac{1}{|W \cap W_{x_i - x_j}|}$ are given by the area of the intersection of W with the translated by $x_i - x_j$ window $W_{x_i - x_j}$.

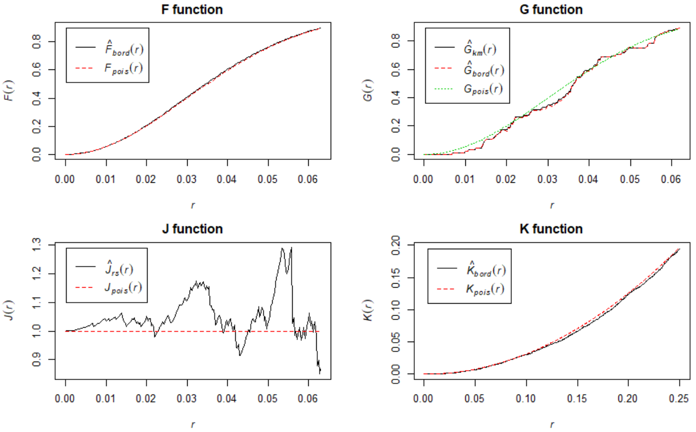


Isotropic edge-correction

- Applicable for stationary and isotropic point processes
- The weights $w(x_i, x_j) = \frac{\nu_1(\partial b(x_i, ||x_i - x_j||) \cap W)}{2\pi ||x_i - x_j||}$ where ν_1 denotes the length of a curve, ∂ denote the boundary of a set and $b(x_i, r)$ the ball centred at x_i with radius r.
- The weights give the proportion of the circle that lies in W.

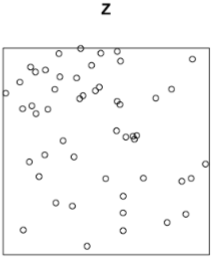


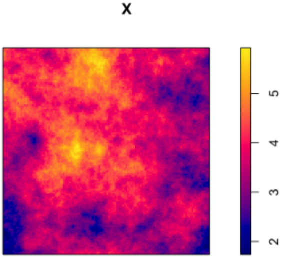
Summary statistics for CSR



Log-Gaussian Cox process

- Hierarchical model, where X is a Gaussian random field and $Z \mid X$ is an inhomogeneous Poisson process where $\lambda(x, y) = exp(X(x, y))$
- Example: X is a Gaussian random field with mean 3 and an exponential covariance function.





Neyman-Scott processes

Cluster processes are models for aggregated spatial point patterns

For Neyman-Scott cluster process

(MC1) parent events form a Poisson process with intensity λ

(MC2) each parent produces a random number S of daughters (offsprings), realized independently and identically for each parent according to some probability distribution p_s

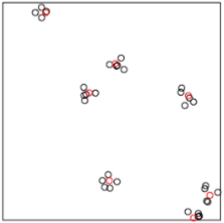
(MC3) the locations of the daughters in a cluster are independently and identically distributed according to a bivariate continuous probability density function.

The cluster process consists only of the daughter points.

Examples of Neyman-Scott processes

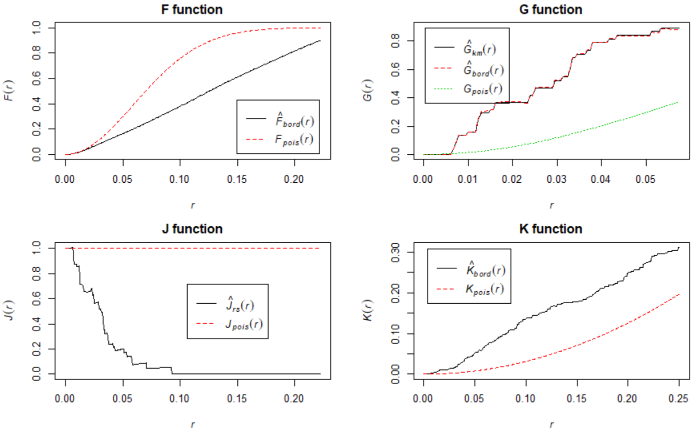
- Matérn cluster process: p_s is a Poisson distribution and the continuous distribution for the locations of offspring is the uniform distribution on a disc
- Thomas cluster process:p_s is a Poisson distribution and the continuous distribution for the locations of offspring is the the 2-dimensional normal distribution

Realization of a Matérn cluster process



Red: Parent points from a Poisson point process with intensity 7
Black: Daughter points with cluster radius 0.05 and average number of daughter per cluster 5.

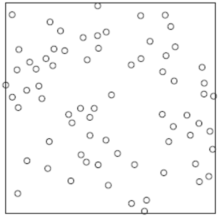
Summary statistics

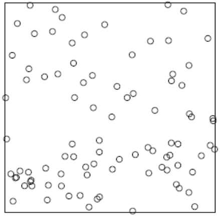


Matérn I hard-core process

- Hard-core processes are models for regular spatial point patterns
- Matérn I hard-core process:
  1. Simulate a homogeneous Poisson process Z
  2. Delete any point in Z that lies closer than a distance r from the nearest other point
- There is a minimum allowed distance, called hard-core distance, between any two points
- Matérn I hard-core process: A Poisson process with intensity λ is thinned by deleting all pairs of points that are at distance less than the hard-core radius apart.

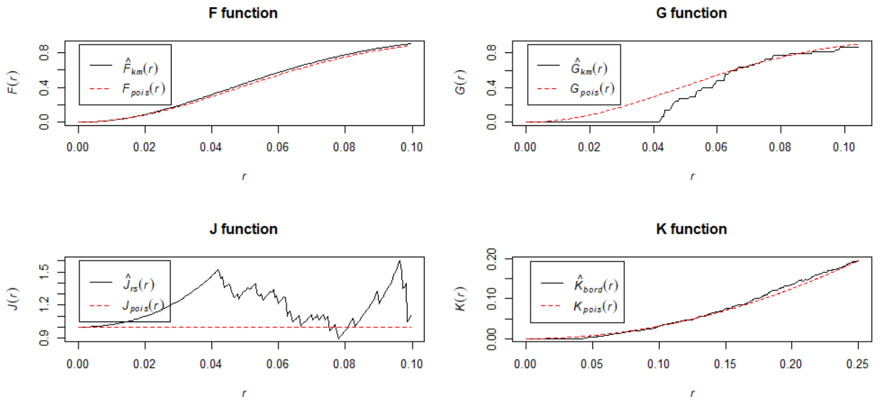
Realization of a Matérn I hard-core process





Left: Hard-core process with the initial Poisson intensity 300, hard-core radius 0.04
Right: Poisson process with intensity 100

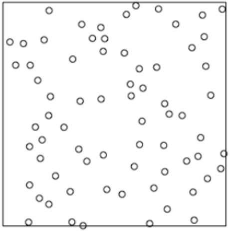
Summary statistics



Matérn II hard-core process

The thinning strategy for Matérn II hard-core processes is the following

1. Simulate a homogeneous Poisson process Z
2. Mark each point in Z by "ages", which are independent and uniformly distributed numbers in $[0,1]$.
3. Delete any point in Z that lies closer than a distance r from another point that has has a higher age.



Pairwise interaction processes

- Pairwise interaction processes are a subclass of Markov point processes which are models for point patterns with interaction between the events
- There is interaction between the events if they are "neighbours", e.g. it they are close enough to each other
- Models for inhibition/regularity

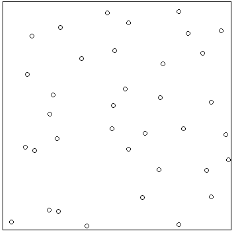
Pairwise interaction processes: Strauss process

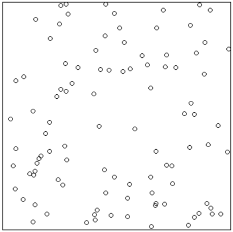
- Two points are neighbours if they are closer than distance R apart
- The density function (with respect to a Poisson process with intensity 1) is

$$f(x) = \alpha\beta^{n(x)}\gamma^{s(x)}, \ \beta > 0, \ \gamma \geq 0,$$

  where
  - $\beta > 0$ is the effect of a single event (connected to the intensity of the process)
  - $0 < \gamma \leq 1$ is an interaction parameter
  - $n(x)$ is the number of points in the configuration
  - $s(x)$ is the number of R close pairs in the configuration, where $R > 0$ is an interaction radius (range of interaction)
  - α is a normalizing constant

Example: Strauss process





Left: Strauss process with $\beta = 100$, $\gamma = 0.2$ and $R = 0.1$
Right: Strauss process with $\beta = 100$, $\gamma = 1$ and $R = 0.1$

- $\gamma = 1$ corresponds to CSR
- $\gamma < 1$ corresponds to inhibition

Simulation based envelope tests

- Let $\hat{K}(r)$ be an estimate of the K function for X.
- Comparing $\hat{K}(r)$ with the theoretical value under CSR ($K(r) = \pi r^2$) we can get information about whether there is clustering or inhibition at scale r.
- But how large deviation could we expect to find by pure randomness?
- Simulation based envelope techniques have been developed to tackle this issue.

Simulation based envelope tests

- Let $N(W) = n$, then we
  1. Simulate $X_1, ..., X_M$ on W from a Poisson process conditioned on $N_{X_i}(W) = n$ (Binomial point process).
  2. Calculate $\hat{K}_i(r)$ for all X_i and check if

$$\min_i \hat{K}_i(r) \leq \hat{K}(r) \leq \max_i \hat{K}_i(r)$$

- For a fixed value of r and for $M = 39$ this is a valid test at level $\alpha = 0.05$.
- However, the probability that the K function will wander outside the envelope for some r is very much higher than 0.05.
- Recently, statistically valid envelope tests have been developed (Myllymäki et al. 2017)

References

- Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H., Hahn, U. (2017). Global envelope tests for spatial processes. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79(2), 381-404.
- Baddeley, A.,Turner, R. Spatstat: an R package for analyzing spatial point patterns. *J. Stat. Softw.* 12 (2005) 1-42.
- Besag, J.E., 1977. Comment on "Modelling spatial patterns" by B. D. Ripley. *Journal of the Royal Statistical Society* B (Methodological) 39 (1977) 193-195.
- Illian, J., Penttinen, A., Stoyan, H., Stoyan, D. *Statistical Analysis and Modelling of Spatial Point Patterns*. Chichester: Wiley (2008).
- Ripley, B.D. Modelling spatial patterns. *Journal of the Royal Statistical Society* B 39 (1977) 172-212.